

Rongxing DUAN  
Yanni LIN  
Yining ZENG

## FAULT DIAGNOSIS FOR COMPLEX SYSTEMS BASED ON RELIABILITY ANALYSIS AND SENSORS DATA CONSIDERING EPISTEMIC UNCERTAINTY

### DIAGNOZOWANIE BŁĘDÓW W SYSTEMACH ZŁOŻONYCH NA PODSTAWIE ANALIZY NIEZAWODNOŚCI ORAZ DANYCH Z CZUJNIKÓW Z UWZGLĘDNIENIEM NIEPEWNOŚCI EPISTEMICZNEJ

*This paper presents an information fusion method to diagnose system fault based on dynamic fault tree (DFT) analysis and dynamic evidential network (DEN). In the proposed method, firstly, it uses a DFT to describe the dynamic fault characteristics and evaluates the failure rate of components using interval numbers to deal with the epistemic uncertainty. Secondly, qualitative analysis of a DFT is to generate the characteristic function via a traditional zero-suppressed binary decision diagram, while quantitative analysis is to calculate some importance measures by mapping a DFT into a DEN. Thirdly, these reliability results are updated according to sensors data and used to design a novel diagnostic algorithm to optimize system diagnosis. Furthermore, a diagnostic decision tree (DDT) is obtained to guide the maintenance workers to recover the system. Finally, the performance of the proposed method is evaluated by applying it to a train-ground wireless communication system. The results of simulation analysis show the feasibility and effectiveness of this methodology.*

**Keywords:** dynamic fault tree, dynamic evidential network, interval numbers, sensors data, diagnostic importance factor.

*W artykule przedstawiono metodę fuzji informacji służącą do diagnozowania błędów systemu w oparciu o analizę dynamicznego drzewa błędów (DFT) oraz dynamiczną sieć dowodową (DEN). W proponowanej metodzie, pierwszym krokiem jest wykorzystanie DFT do opisanie dynamicznych charakterystyk błędów oraz ocena intensywności uszkodzeń komponentów przy użyciu liczb przedziałowych, która rozwiązuje problem niepewności epistemicznej. Krok drugi stanowi jakościowa analiza DFT, która polega na wygenerowaniu funkcji charakterystycznej za pomocą tradycyjnego binarnego diagramu decyzyjnego typu "zero-suppressed" (w którym zostały wyeliminowane wszystkie węzły, których krawędź „1” prowadzi do liścia „0”), oraz analiza ilościowa polegająca na obliczeniu pewnych miar ważności poprzez odwzorowanie DFT w DEN. W kroku trzecim, otrzymane wyniki niezawodnościowe aktualizuje się zgodnie z danymi z czujników a następnie wykorzystuje do stworzenia nowego algorytmu diagnostycznego do optymalizacji diagnostyki systemu. Powstaje diagnostyczne drzewo decyzyjne (DDT), które stanowi dla pracowników utrzymania ruchu wytyczną w procesie odzyskiwania systemu. Działanie proponowanej metody oceniano poprzez zastosowanie jej do diagnostyki systemu łączności radiowej pociąg–ziemia. Wyniki analizy symulacyjnej wskazują na możliwość praktycznego wykorzystania i skuteczność omawianej metodologii.*

**Słowa kluczowe:** dynamiczne drzewo błędów, dynamiczna sieć dowodowa, liczby przedziałowe, dane z czujników, czynnik ważności diagnostycznej.

#### 1. Introduction

With the rapid development of science and technology, application of high dependability safeguard techniques have improved the performance of modern systems greatly on the one hand, but increased the complexity of these systems on the other hand, which significantly raises some challenges in fault diagnosis. These challenges are failure dependency of components and epistemic uncertainty. Usually, some methods of fault tolerance are used to improve the system reliability. The behaviours of components in this system, such as failure priority, functional dependent failures, and sequentially dependent failures should be taken into account. In addition, high reliability makes it extremely difficult to obtain complete fault data because these systems may still be in the early life cycle, which results in the epistemic uncertainty. Thus, the work of fault diagnosis has attracted more attention than before. The aim of a fault diagnosis system is to quickly detect and identify the root causes of these failures based on some in-

formation such as sensors data and operator experience by using some models and algorithms. Several efficient fault diagnosis approaches have been proposed for a variety of systems over the last few decades. Doguc et al. proposed a new fault diagnosis method based on the real-time reliability analysis [7]. Bayesian network (BN) was used to calculate the system reliability, and the real-time system reliability was monitored and compared with the previous values. If the deviations exceeded the set threshold, a heuristic efficient algorithm was used to locate the failed component which had the greatest changes between the prior probability and posterior probability. In the literature [3], a real-time fault diagnosis method for complex systems using object-oriented BN was proposed. It included an off-line BN construction phase and an on-line fault diagnosis phase. However, the construction of BN model requires a large amount of fault data. In [5], a fault diagnosis approach based on the fuzzy neural network and fault tree was proposed. Fuzzy neural network was used to train the relation-

ship between faults and symptoms. Fault tree was used to describe the logical relationship between faults and symptoms. In [13], a new method was proposed to diagnose the bearing fault using evidence network and support vector machine. The fault model construction was established using a data-driven method, and the evidence theory was used to solve the conflicting results from different layer models to increase the diagnosis accuracy. However, the above methods are based on the data-driven fault method which needed lots of fault data and cannot deal with the epistemic uncertainty. A fault diagnosis method for safety instrumentation system based on the fault tree and BN was proposed [6]. It used the static fault tree to construct the fault model of safety instrument system and mapped the fault tree into BN to calculate the importance measure which was used to design the diagnosis algorithm. Nevertheless, this method is unable to describe the dynamic fault characteristics and fails to deal with the epistemic uncertainty. In work of [1], DFT was introduced to model the dynamic fault behaviours and diagnostic importance factor (DIF) was calculated to determine the diagnostic sequence. However, this method determined the diagnosis sequence only by components' DIF, and usually caused minimal cut sets ((MCS)) with a smaller DIF to be checked first, thereby influencing the diagnosis result. Tao et al. presented an improved fault diagnosis method which took components' DIF and MCS's DIF into account to avoid that case [23]. In order to improve the diagnosis efficiency, Assaf et al. proposed a method to incorporate the evidence information from sensors into the diagnostic process based on the DFT [2]. However, the solution for DFT is based on Markov chains, which is ineffective in handling large DFT and modelling power capabilities. Furthermore, it cannot update the reliability results according to the evidence data from sensors, which affects the diagnostic efficiency. Therefore, Duan et al. presented an efficient diagnostic algorithm which used DFT to establish a system failure model and calculated reliability parameters using a discrete time Bayesian network (DTBN) [8]. This approach not only can avoid the state space explosion, but also can incorporate sensor information to update reliability results. Nevertheless, DTBN is an approximate method to solve DFT and there is a contradiction between the accuracy and computational complexity. Furthermore, these diagnosis methods are usually assumed that the failure rates of the components are expressed in crisp values describing their reliability characteristics and cannot cope with the epistemic uncertainty. So, a fuzzy DFT analysis was introduced, which can deal with the uncertainty and model the dynamic fault characteristics [12, 17]. Nevertheless, the solution for the fuzzy DFT was still based on the Markov chains. To overcome these shortcomings, a new fault diagnosis algorithm based on fuzzy set and DFT analysis was proposed [10]. The fuzzy information obtained by fuzzy set theory and domain expert was transformed into quantitative information to obtain the fuzzy failure rates of components. DTBN was used for quantitative analysis. Nevertheless, it is usually difficult to determine the corresponding membership function of each language value. To this end, Duan et al. proposed a new fault diagnosis for complex systems based on dynamic evidential network and multi-attribute decision making [11]. It used interval numbers to express the failure rates of the basic events and obtained the optimal diagnosis sequences based multi-attribute decision making with interval numbers. However, this method failed to incorporate the sensors data to optimize the diagnosis process.

In summary, fault diagnosis methods based on reliability analysis have some following limitations:

- (1) Traditional fault diagnosis methods based on reliability analysis generally use a static fault tree or DFT to construct fault model and assume that the failure rates of all events are crisp values, which cannot deal with epistemic uncertainty. Although some researchers put forward the possibility theory [21, 25], fuzzy set theory [4, 15], imprecise probability [18], interval analysis [27] and evidence theory [28], these theories were only used for

the reliability analysis and risk assessment and were not further applied to the fault diagnosis. Furthermore, Markov chains and DTBN are usually used to solve DFT. Markov chains have a bad state space explosion problem and the inability to update the posterior probability of the component based on sensors data. The DTBN based solution for DFT has the contradiction between computational accuracy and computational complexity. That is, its computational accuracy is related to the size of time granularity  $n$ . As  $n$  increases, the conditional probability table has an exponential growth [26]. Although the solution proposed in [16] can solve the problem of calculation accuracy to a certain extent, it cannot fuse the sensors information for backward reasoning.

- (2) From the aspect of sensors information fusion, Traditional method appends a sensor layer for capturing evidence onto the DFT without impacting the reliability analysis, and the sensor layer uses static gates to represent evidence information. However, evidence information is only used to update qualitative information to reduce the number of suspected MCS and fails to update the quantitative information, thus unable to reflect the contribution of components to the system failure.
- (3) In the view of the diagnosis algorithm, the algorithms based on reliability analysis generally only take the importance measures or posterior probability of components into account [1, 9]. Furthermore, the importance measures are usually crisp values and cannot be used to make decisions under uncertainty.

Motivated by the problems motioned above, this paper proposes an information fusion method to diagnose system fault based on DFT and DEN. DFT is used to establish the system fault model to describe the dynamic fault characteristics. Interval numbers are used to describe the failure rate of components to deal with epistemic uncertainty. Furthermore, an efficient zero-suppressed binary decisions diagrams is used to obtain all MCSs, and a DFT is mapped into a DEN to calculate the reliability parameters. In addition, evidence information from sensors is incorporated to update the qualitative information and quantitative parameters, which are used to design the fault diagnosis algorithm. Finally, a train-ground wireless communication system is given to demonstrate the efficiency of this proposed method.

The remainder of this article is organized as follows. Section 2 presents the model construction and qualitative analysis of DFT. Section 3 introduces the dynamic evidence network and provides a quantitative analysis method by mapping a DFT to a DEN. A novel approach is proposed to incorporate the evidence information to update the reliability results, and an efficient diagnosis algorithm is given in Section 4. Section 5 is devoted to a simple illustration example of the proposed approach. Some conclusions and future research recommendations are given in the final section.

## 2. DFT

### 2.1. Model Construction of DFT

Fault tree is a deductive method to decide the potential causes that may cause the occurrence of a predefined undesired event, generally denoted as the top event. DFT extends a static fault tree to describe the dynamic failure behaviours such as priorities of failure events, spares, and sequence-dependent events. Dynamic gates in DFT include the priority AND gate (PAND), the functional dependency gate (FDEP), the sequence enforcing gate (SEQ), the cold, hot, and warm spare gates (CSP, HSP, WSP). The model construction of the fault tree usually requires an in depth knowledge of the system and its components. It includes the construction of a network topology and the failure rates estimation of components. The former can resort to fault mode and effect analysis and the latter needs to obtain lots of fault data, which

is almost impossible to estimate precisely the failure rates of the basic events in the practical engineering application. In this paper, interval numbers are used to describe the failure rates of the basic events based on the expert elicitation and some data sheet at the design stage.

**2.2. Qualitative analysis of a DFT**

The qualitative analysis of a fault tree can be used to obtain the MCS. Algebraic simplification is the most effective method to solve MCS, but it is not suitable for solving DFT. Zero-suppressed binary decisions diagrams, introduced by Tang, separate timing constraints and logic constraints and convert a DFT into a static fault tree [24]. This algorithm generates the MCS of the corresponding static fault tree using several set operations and then it can be expanded into minimal cut sequences if we consider the timing constraints.

Let  $S_1, S_2$  be the input of MCS for AND gate and MCS for OR gate respectively, several set operations are as follows:

$$\begin{aligned} S_c &= S_1 \cap S_2, D_1 = S_1 - S_c, D_2 = S_2 - S_c \\ U &= D_1 \cup D_2, P = D_1 * D_2, D_3 = U - P \\ MCS_{OR} &= S_c \cup D_3, MCS_{AND} = S_c \cup P \end{aligned} \tag{1}$$

where  $D, S_c, U,$  and  $P$  respectively represent set difference, set intersection, set union, and set product. MCSOR and MCSAND are the output of MCS- OR and MCS- AND respectively.

The MCS generation algorithm is implemented recursively during the depth-first left-most traversal of a fault tree. Firstly, it generates the MCS of the inputs of a connection gate, and then executes some operations to combine the MCS of the inputs into the MCS of the output of the connection gate. Finally, all the minimal cut sequences from the MCS can be obtained by considering the timing constraints [24].

**2.3. Quantitative analysis of a DFT**

Quantitative analysis of a DFT is mainly to calculate the system reliability and some importance measures. DIF is the most frequently used importance measure and is also the cornerstone of diagnosis method based on reliability. From a diagnostic point of view, it allows us to discriminate between components by their importance. It is well known to us all that components with a larger DIF value should be diagnosed first. It can assure a minimal number of system checks while bringing back the system. Reliability parameters are calculated by converting a DFT into a DEN which is introduced in Section 3.

**3. DEN**

**3.1. EN**

D-S evidence theory has a unique ability in the expression of epistemic uncertainties. The evidence theory can be well compatible with the theory of probability. EN consists of BN and D-S evidence theory and includes both advantages [14]. It is a popular analysis tool for representing and managing epistemic uncertainties. An EN is a directed acyclic graph (DAG) used to represent system's uncertain knowledge and system logic in artificial intelligence. An EN is defined as  $EN = \langle G, P \rangle$ , where  $G = \langle N, A \rangle$  represents a network graph and  $N = \{N_1, N_2, \dots, N_k\}$  represents a set of nodes. A node can be a basic variable or an abstraction of a system or component, such as system reliability, component status.  $A$  is a set of arcs, which indicate direct conditional relations between the connected nodes.  $P$  represents some network parameters in EN. Each network parameter represents the belief distributions that are distributed to a node, and each node  $X_i \in N$  has a corresponding conditional belief table. The parent node

of node  $X_i$  is set to  $Pa(X_i)$  and their relationship is expressed in the formula  $P(X_i | Pa(X_i))$ .

**3.2. DEN**

A DEN extends an EN with adding a temporal dimension. This new dimension is managed by defining different nodes to model variables with respect to different time slices. A DEN includes an initial network and some temporal transition networks. Each time slice corresponds to a static EN, and the time slices are made up of a directed acyclic graph  $G_T = \langle V_T, E_T \rangle$  and the corresponding conditional probabilities. The  $V_T$  and  $E_T$  are respectively nodes of time  $T$  and directed arcs. A directed arc links two variables belonging to different time slices and  $E_T^{tmp}$  is used to denote the temporal transition network of time slices. Then  $E_T^{tmp}$  can be determined by:

$$E_T^{tmp} = \{(a,b) | a \in V_{T-1}, b \in V_T\}, T_0 \leq T \leq T_0 + N\Delta T \tag{2}$$

where  $T_0$  is an initial network.

In the DEN model,  $G_T$  depends solely upon the present state and the previous state. Thus, the following equation is obtained:

$$P(G_T | G_{T-\Delta T}, \dots, G_{T_0}) = P(G_T | G_{T-\Delta T}) \tag{3}$$

In addition, we define these impacts as transition-belief masses between the focal elements of the variable at time step  $k$  and those at time step  $k+1$  and the CBT relative to inter-time slices is calculated by Equation 4:

$$m(X_{k+1} | X_k) = \begin{bmatrix} m(G_1^{X_{k+1}} | G_1^{X_k}) & \dots & m(G_Q^{X_{k+1}} | G_1^{X_k}) \\ \vdots & \ddots & \vdots \\ m(G_1^{X_{k+1}} | G_Q^{X_k}) & \dots & m(G_Q^{X_{k+1}} | G_Q^{X_k}) \end{bmatrix} \tag{4}$$

**3.3. System reliability model of DEN**

In evidence theory,  $\Theta = \{W_i, F_i\}$  is the knowledge framework of the component  $i$  and the focal elements are defined by:

$$2^\Theta = \{\{\emptyset\}, \{W_i\}, \{F_i\}, \{W_i, F_i\}\} \tag{5}$$

where  $\{W_i\}$  and  $\{F_i\}$  denote the working state and failure state respectively. The state of  $\{W_i, F_i\}$  corresponds to the epistemic uncertainty.

Belief measure (*Bel*) defines the lower bound of the probabilities that the focal element exists, and plausibility measure (*Pl*) defines the upper bound of the probabilities that the focal element exists. The basic belief assignment on the system state expresses an epistemic uncertainty, where *Bel* and *Pl* measures are not equal and bound the system reliability. Therefore, the basic probability assignment (BPA) of component  $i$  can be computed as:

$$\begin{cases} m(\{W_i\}) = Bel(\{W_i\}) \\ m(\{F_i\}) = 1 - Pl(\{W_i\}) \\ m(\{W_i, F_i\}) = Pl(\{W_i\}) - Bel(\{F_i\}) \end{cases} \tag{6}$$

Presumably, the upper and lower bounds of the component's failure probability is equivalent to the BPA in the DEN:

$$\begin{cases} m(\{W_i\}) = 1 - \overline{P(x)} \\ m(\{F_i\}) = \underline{P(x)} \\ m(\{W_i, F_i\}) = \overline{P(x)} - \underline{P(x)} \end{cases} \quad (7)$$

where  $Bel(\{F_i\}) = \underline{P(x)}$  and  $Pl(\{F_i\}) = \overline{P(x)}$ .

### 3.4. DFT analysis based on DEN

#### 3.4.1. Converting a static logic gate into a DEN

Static logic gates mainly include three gates, AND gate, OR gate and voting gate. This section takes an OR gate for example and provides the schemes to map an OR gate into a DEN. When any of the input components  $X_i$  ( $i=1, \dots, n$ ) of an OR gate fails, the output of the gate fails too. Fig. 1 shows an OR gate and the equivalent DEN. Table 1 gives the conditional probabilities of node  $A(T+\Delta T)$  in the DEN. Equation 8 gives the conditional probabilities of output node  $E(T+\Delta T)$ . A more detailed description of this work can be found in [20].

Table 1. The conditional probabilities of node  $A(T+\Delta T)$ .

$A(T)$	$A(T+\Delta T)$		
	{W}	{F}	{W,F}
{W}	$m_A(W)$	$m_A(F)$	$m_A(W,F)$
{F}	0	1	0
{W,F}	0	$m_A(F)$	$1 - m_A(F)$

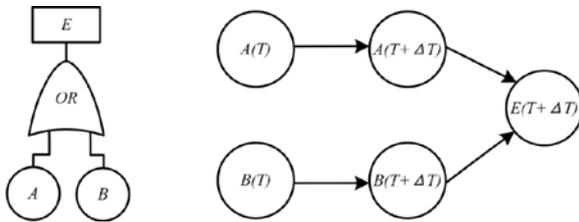


Fig. 1. An OR gate and the equivalent DEN

$$\begin{cases} P(E=1 | A(T+\Delta T)=0, B(T+\Delta T)=1) = 1 \\ P(E=1 | A(T+\Delta T)=1, B(T+\Delta T)=0) = 1 \\ P(E=1 | A(T+\Delta T)=1, B(T+\Delta T)=1) = 1 \\ P(E=1 | A(T+\Delta T)=1, B(T+\Delta T)=\{0,1\}) = 1 \\ P(E=1 | A(T+\Delta T)=\{0,1\}, B(T+\Delta T)=1) = 1 \\ P(E=\{0,1\} | A(T+\Delta T)=0, B(T+\Delta T)=\{0,1\}) = 1 \\ P(E=\{0,1\} | A(T+\Delta T)=\{0,1\}, B(T+\Delta T)=0) = 1 \\ P(E=\{0,1\} | A(T+\Delta T)=\{0,1\}, B(T+\Delta T)=\{0,1\}) = 1 \\ P(E=1 | A(T+\Delta T)=0, B(T+\Delta T)=0) = 0 \end{cases} \quad (8)$$

#### 3.4.2. Converting a dynamic logic gate into a DEN

Some dynamic logic gates are introduced to model the functional and sequential in the DFT. These logic gates include PAND, SEQ, FDEP and spare gates. An FDEP gate will be used to describe how the dynamic logic gates are mapped into DEN. An FDEP gate includes a trigger event and some dependent basic events. The trigger event can be a basic event or an output of another gate in the DFT. The occurrence of a trigger event will force all basic events to occur, which

means all basic events functionally depend upon the trigger event. Fig. 2 shows an FDEP gate and the equivalent DEN. Table 2 and Table 3 show the conditional probabilities of the node  $A(T+\Delta T)$  and  $E(T+\Delta T)$  respectively.

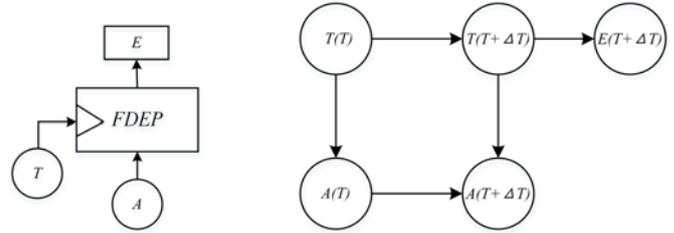


Fig. 2. An FDEP gate and the equivalent DEN

Table 2. The conditional probabilities of the node  $A(T+\Delta T)$

$T(T+\Delta T)$	$A(T)$	$A(T+\Delta T)$		
		{W}	{F}	{W,F}
{W}	{W}	$m_A(W)$	$m_A(F)$	$m_A(W,F)$
{W}	{F}	0	1	0
{W}	{W,F}	0	0	1
{F}	{W}	0	1	0
{F}	{F}	0	1	0
{F}	{W,F}	0	1	0
{W,F}	{W}	0	0	1
{W,F}	{F}	0	1	0
{W,F}	{W,F}	0	0	1

Table 3. The conditional probabilities of the node  $E(T+\Delta T)$ .

$T(T+\Delta T)$	$E(T+\Delta T)$		
	{W}	{F}	{W,F}
{W}	1	0	0
{F}	0	1	0
{W,F}	0	0	1

#### 3.4.3. Calculating reliability results

After DFT model of a system is built, it can be mapped into the equivalent DEN using the approach mentioned above. Reliability results of system can be obtained by resorting to the DEN inference algorithm. Reliability parameters mainly include system unreliability and DIF, which can be used to develop a diagnosis algorithm.

The unreliability of a system is calculated by the following equation:

$$P_S = [P_S, \overline{P_S}] = [Bel(\{F_S\}), Pl(\{F_S\})] \quad (9)$$

where  $[Bel(\{F_S\}), Pl(\{F_S\})]$  represents the failure probability of a system.

DIF is usually defined as the probability that a basic event has occurred given that the top event has also occurred. The DIF of a component  $i$  is given by:

$$DIF_i = P(i | S) = [Bel(\{F_{i|S}\}), Pl(\{F_{i|S}\})] \quad (10)$$

where  $i$  is a component in the system  $S$ ;  $P(i|S)$  is the probability that the basic event  $i$  has occurred given the top event has occurred.

Similarly, the DIF of a  $MCS_n$  is defined by:

$$DIF_{MCS_n} = P(MCS_n|S) = \frac{P(MCS_n)}{P(S)} \quad (11)$$

where  $P(S)$  is the unreliability of the system  $S$ ;  $P(MCS_n|S)$  is the failure probability that the  $MCS_n$  has occurred given the top event has occurred.

For convenience, we calculate the value  $P(MCS_n)$  instead of  $DIF_{MCS_n}$  and use it to design the diagnosis algorithm in the following section.

### 3.5. Importance sorting using possibility-based NSG ranking approach

Based on above analysis, we can obtain the interval value of DIF which can be used to develop an efficient diagnosis algorithm in order to reduce the diagnosis cost. As is known to all, components with a larger DIF are more important from a diagnostic point of view. Thus, the importance ranking of components will be very important for determining a diagnosis sequence. Nevertheless, these interval values are not sufficient to rank components and should be converted into a probability measure. In this paper, a possibility-based NSG ranking method, developed by Nakahara et al. is used to rank DIF of components expressed by interval numbers [19, 22]. This method can be used to compare the DIF of components to provide a guidance for system diagnosis.

For interval numbers  $a = [a^-, a^+]$  and  $b = [b^-, b^+]$ ,  $l(a)$  and  $l(b)$  respectively denote the lengths of the intervals  $a = [a^-, a^+]$  and  $b = [b^-, b^+]$ , it calculated as follows:

$$l(a) = a^+ - a^-, l(b) = b^+ - b^- \quad (12)$$

Then the possibility of  $[a] \geq [b]$  can be defined as:

$$p([a] \geq [b]) = \min\{0, 1 - \max(\frac{a^+ - b^-}{l(a) + l(b)}, 0)\} \\ = \begin{cases} 1 & a^- \geq b^+ \\ \frac{a^+ - b^-}{l(a) + l(b)} & a^+ > b^- \text{ and } a^- < b^+ \\ 0 & a^+ \leq b^- \end{cases} \quad (13)$$

A possibility-based NSG ranking method includes the following steps.

**Step 1:** For a set of interval numbers  $a_i = [a_i^-, a_i^+]$   $i=1,2,\dots,n$ , compare them with each other, and then the corresponding possibility  $p_{ij} = p(a_i > a_j)$  can be obtained. So we can establish the probability matrix  $P = (p_{ij})_{n \times n}$ , which is given by:

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \quad (14)$$

**Step 2:** Denote  $\lambda_i = \sum_{j=1}^n p_{ij}$  as the row sum of the possibility matrix  $P$  and  $\lambda = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_n)^T$  as the corresponding row sum vector.

**Step 3:** Calculate the ranking vector  $\omega = (\omega_i)$  is given by:

$$\omega_i = \frac{1}{n(n-1)} (\lambda_i + \frac{n}{2} - 1) \quad i = 1, 2, \dots, n \quad (15)$$

According to the ranking vector of the possibility matrix  $P$ , the interval numbers  $a_i = [a_i^-, a_i^+]$   $i=1,2,\dots,n$  can be sorted based on the value of  $\omega_i$ .

## 4. Fault Diagnosis Method based on Reliability Analysis and Sensors data

### 4.1. Model construction of diagnostic sensors

When a system fails, usually several evidence information from sensors can be observed too, and this may be utilized to improve the efficiency of the diagnosis algorithm. In general, the more the number of sensors used to monitor the system, the higher the diagnostic efficiency of the system. However, too many sensors will increase system costs on the one hand, but on the other hand, it will reduce the reliability of the diagnostic system. So a tradeoff between the good points against the bad should be taken into account. Besides, sensors might fail and false information can misguide the diagnosis process. For simplicity, we assume that sensors never fail in the paper. To optimize the diagnosis process, a diagnostic sensors model is constructed to update the qualitative and quantitative information. As we all know, the DEN created from DFT has no evidence nodes representing the evidence information, thus, we need to add them in the DEN. Evidence nodes in the DEN provide links connecting it with the component in the DEN, which are monitored by sensors. The links are directed from the component to the evidence nodes. Evidence nodes in the DEN create a conditional probability table using the probability of producing the observation results. This diagnostic sensors model does not affect the system reliability analysis and can update the qualitative information and quantitative parameters according to sensors data.

### 4.2. Incorporating sensors data

#### 4.2.1. Updating the system characteristic function

If sensors detect some failed components, we can use this evidence information to minimize the number of the diagnosed MCS. Since, examining a cut set that caused the system to fail then fixing the failed components in that cut set should recover the system, we can increase the efficiency of fault diagnosis by reducing the number of cut sets examined. The cut sets under evidence (CUE) is the set of all essential MCS obtained after evidence information removes some unsuspected cut sets. We can use evidence information from sensors to simplify the characteristic function of the system in order to obtain the CUE function using the algorithm in [2].

#### 4.2.2. Updating DIF

In addition, we can use the evidence information from sensors to update DIF, which reflects objectively the contribution to the system failure. The DIF of the components under the evidence information conditions can be calculated using the Equation (16). Calculating DIF is very simple. We just input the corresponding evidence information

to the DEN and obtain the DIF of components and CUE using the inference algorithm:

$$DIF'_i = P(i|S, E) = \frac{P(i, E, S)}{P(S)DIF_E} \quad (16)$$

where  $i$ ,  $S$  and  $E$  represent a component, system and evidence information, respectively.

**4.3. Fault diagnosis algorithm**

The aim of fault diagnosis is to obtain the optimal check sequence to locate the fault as fast as possible using an efficient diagnosis algorithm. As it is known to all, the direct cause of the system failure is the failure of a CUE. So, we should check CUE one by one to locate the failed component in the system. Only when we finish checking a CUE can we do next. The sequence by which CUE is diagnosed depends on the corresponding DIF, while the sequence of components in the same CUE is determined by their DIF. The CUE with a larger DIF is checked first. Accordingly, the component with a larger DIF in a CUE is checked first. It can assure a minimal number of system checks while bringing the system back. The fault diagnosis algorithm, which incorporates sensors data, is as follows:

**Step 1.** List all CUEs and rank them according to their DIF.

**Step 2.** Select the CUE with a highest DIF value and diagnose the component  $X$  with a highest DIF in the same CUE.

**Step 3.** Split all CUEs into those with  $X$  and those without.

a) If  $X$  has failed test, we take all CUEs that include  $X$

- Diagnose all CUEs and the CUE with a higher DIF is checked first.
- The component with a larger DIF in the same CUE is checked first.

b) If  $X$  has not failed test, we take the other CUEs

- Select the CUE untested with a highest DIF value.
- And recursively repeat Step2 - Step3.

**4.4. Evaluation of diagnosis algorithm**

The diagnosis algorithm can easily be described in the graphical DDT, which can help us recognize the failed components with a map. It is a directed acyclic graph composed of circular nodes and arcs linking parent nodes to child nodes. A node represents a component being tested. Arcs point to the next component to be tested; right arcs point to components within the same cutset as the parent node, and left arcs point to components which are not in the same cutset as the parent node. Moreover, when diagnostician reaches a node and tests the component at the node, the test either fails or passes. If the test fails, then the right arc is traversed indicating the need to repair the tested component in the parent node. If a test passes, then the left arc is traversed indicating that the cut sets which include the tested component in the parent node have not failed.

There are many indicators to evaluate the fault diagnosis algorithm. In this paper, we can evaluate the diagnostic efficiency with the help of the DDT. Traditional evaluation measures only take the test cost or the failure probability

of components into account, and neglect the qualitative information and the importance factors. Thus, we use expected diagnostic cost (EDC) which incorporates the structure information, DIF and test cost into one measure for predicting diagnosis cost. This evaluation index takes the diagnosis accuracy as well as the diagnosis cost into account and also considers the relationship between component failure and system failure. Generally, the diagnostic cost is lower, the diagnostic approach is more efficient. EDC can be computed by:

$$EDC = \sum_{i=1}^n DIF_{CUE_i} cp_i = \sum_{j=1}^{m_i} t_{c_j} \quad (17)$$

where  $DIF_{CUE_i}$  is the DIF of the  $i^{th}$  CUE;  $cp_i$  is the sum of all test cost from the top node to the  $i^{th}$  CUE's leaf node;  $t_{c_j}$  is the test cost of the node  $c_j$ .

**5. A numerical example**

Train-ground wireless communication system is a key subsystem of urban rail transit, and its reliability has been improved by the application of high technologies to ensure safe operation. Once breaking down, less causes the operation performance drop, more leads to a disaster. Therefore, an efficient diagnosis strategy should be taken to restore normal operation as soon as possible. A DFT model of a train-ground wireless communication system is shown in Fig.3. It is assumed that all components have the exponential distribution and interval failure rates of components expressed in interval values are shown in Table 4.

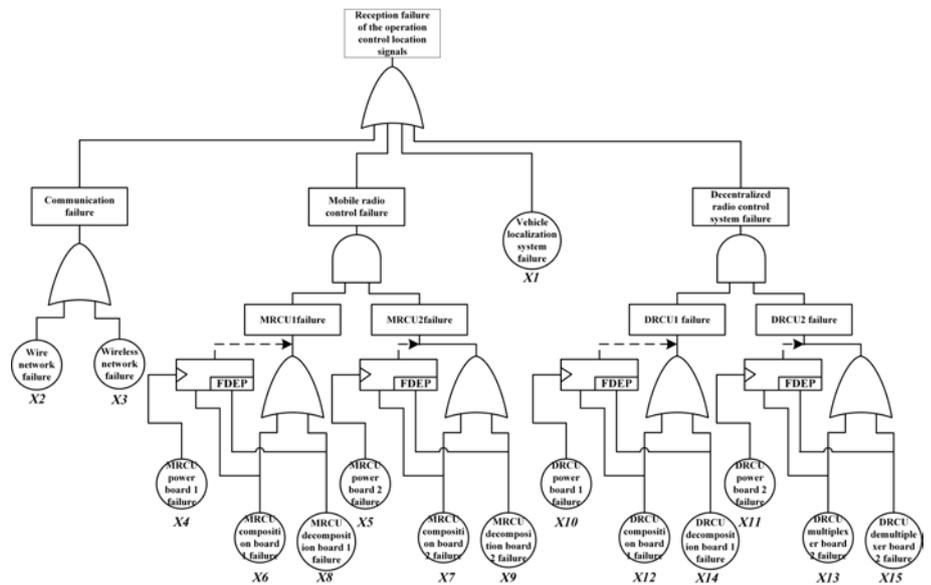


Fig. 3. DFT model of train-ground wireless communication system

Table 4. Failure rates of components are expressed in interval numbers

Components	Interval failure rates	Components	Interval failure rates
X1	[4.22e-6, 5.28e-6]	X8,X9	[5.49e-6, 6.71e-6]
X2	[5.94e-6, 7.26e-6]	X10,X11	[3.15e-5, 3.85e-5]
X3	[4.86e-5, 5.94e-5]	X12,X13	[6.12e-5, 7.48e-5]
X4,X5	[3.78e-5, 4.62e-5]	X14	[5.04e-5, 6.11e-5]
X6,X7	[6.48e-5, 7.92e-5]	X15	[5.04e-5, 6.11e-5]

Through the qualitative analysis of DFT mentioned above, the system characteristic function (the sum of all MCS) of train-ground wireless communication system is obtained:

$$F = X1 + X2 + X3 + X4X5 + X4X7 + X4X9 + X6X5 + X6X7 + X6X9 + X8X5 + X8X7 + X8X9 + X10X11 + X10X13 + X10X15 + X12X11 + X12X13 + X12X15 + X14X11 + X14X13 + X14X15$$

The DFT is mapped into a corresponding DEN for quantitative analysis. Assuming the task time  $T = 1000$  h, the probability of system failure can be obtained using the inference algorithm and it is [0.08293, 0.10714]. In addition, the DIF of all components and MCSs can be calculated shown in Table 5 and Table 6 respectively.

Table 5. DIFs of all components

Components	DIF of components	Components	DIF of components
X1	[0.0508,0.0518]	X8,X9	[0.0857,0.0939]
X2	[0.0709,0.0722]	X10,X11	[0.0708,0.0756]
X3	[0.5681,0.5727]	X12,X13	[0.2012,0.2156]
X4,X5	[0.0751,0.0822]	X14	[0.1788,0.1914]
X6,X7	[0.1963,0.2148]	X15	[0.1788,0.1914]

Table 6. DIFs of all MCSs

MCSs	DIF of MCSs	MCSs	DIF of MCSs	MCSs	DIF of MCSs
X1	[0.0393,0.0635]	X6.X7	[0.0877,0.1675]	X10.X15	[0.0227,0.0432]
X2	[0.0553,0.0873]	X6.X9	[0.0384,0.0732]	X12.X11	[0.0255,0.0488]
X3	[0.4428,0.6954]	X8.X5	[0.0147,0.0281]	X12.X13	[0.0725,0.1383]
X4.X5	[0.0129,0.0246]	X8.X7	[0.0384,0.0732]	X12.X15	[0.0644,0.1224]
X4.X7	[0.0336,0.0642]	X8.X9	[0.0167,0.0321]	X14.X11	[0.0227,0.0432]
X4.X9	[0.0147,0.0281]	X10.X11	[0.0090,0.0172]	X14.X13	[0.0644,0.1224]
X6.X5	[0.0336,0.0642]	X10.X13	[0.0255,0.0488]	X14.X15	[0.0572,0.1084]

A possibility-based NSG sorting method is used to rank the DIF of components and the ranking vectors  $\omega_i$  of matrices  $P$  can be computed as:

$$\omega_i = (0.0333, 0.0402, 0.1, 0.0544, 0.0544, 0.0873, 0.0873, 0.0643, 0.0643, 0.0446, 0.0446, 0.0889, 0.0889, 0.0738, 0.0738)$$

So, the order of the components' DIF is obtained:

$$X3 > X12(X13) > X6(X7) > X14(X15) > X8(X9) > X4(X5) > X10(X11) > X2 > X1$$

Similarly, the ranking of all MCSs can also be obtained:

$$X3 > X6.X7 > X12.X13 > X12.X15(X13.X14) > X14.X15 > X2 > X6.X9(X8.X7) > X1 > X4.X7(X6.X5) > X10.X13(X12.X11) > X10.X15(X14.X11) > X8.X9 > X8.X5(X4.X9) > X4.X5 > X10.X11$$

We assume that a sensor monitors X6 and detects that it is in a work state. We can use this evidence information to simplify the characteristic function and obtain an updated system characteristic function:

$$F_{CUE} = X1 + X2 + X3 + X4X5 + X4X7 + X4X9 + X8X5 + X8X7 + X8X9 + X10X11 + X10X13 + X10X15 + X12X11 + X12X13 + X12X15 + X14X11 + X14X13 + X14X15$$

In addition, this evidence information can be input into the DEN and the corresponding evidence is as follows:

$$P(X6 = \{W\}) = 1, P(X6 = \{W, F\}) = P(X6 = \{F\}) = 0 \quad (18)$$

Using the DEN reasoning algorithm, the updating DIFs of components and CUEs are shown in Table 7 and Table 8 respectively.

Using the sorting method, we can get the order of components:

$$X3 > X12(X13) > X14(X15) > X7 > X2 > X10(X11) > X1 > X9 > X5 > X8 > X4$$

Based on the proposed diagnosis algorithm, we can get the DDT of train-ground wireless communication system without sensors information, shown in Fig. 4 and the corresponding DDT which incorporates sensors information into diagnosis process shown in Fig. 5.

Since the failure probability of CUE is expressed as an interval number, it cannot be directly used to calculate EDC. For convenience, assuming that all components have a unit test cost and test cost of components is independent, we calculate EDC using the median of the interval number in Equation (17). Table 9 shows the EDC of different diagnostic algorithms and indicates the proposed method is more efficient than others.

#### 4. Conclusion

In this paper, a novel fault diagnosis approach for complex systems is presented based on DFT analysis and DEN, which aims to deal with two important issues that arise in engineering applications, such as failure dependency and epistemic uncertainty. For the challenge of failure dependency, a DFT is used to describe the dynamic fault behaviours. For the challenge of the epistemic uncertainty, the failure rates of components in complex systems are

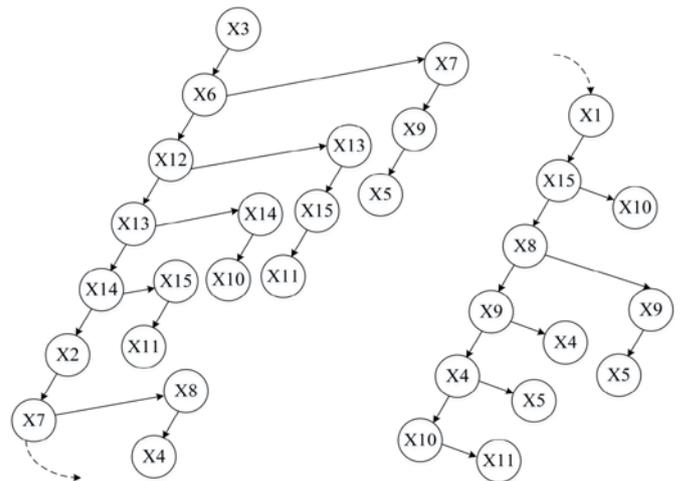


Fig. 4. A DDT of train-ground wireless communication system without sensors information.

Table 7. The updating DIFs of components

Components	DIF of components	Components	DIF of components
X1	[0.0571, 0.0581]	X8	[0.0125, 0.0137]
X2	[0.0797, 0.0809]	X9	[0.0449, 0.0541]
X3	[0.6381, 0.6420]	X10,X11	[0.0762, 0.0807]
X4	[0, 0]	X12,X13	[0.2165, 0.2302]
X5	[0.0394, 0.0474]	X14	[0.1923, 0.2044]
X7	[0.1030, 0.1237]	X15	[0.1923, 0.2044]

Table 8. The updating DIFs of CUEs

CUEs	DIF of CUEs	CUEs	DIF of CUEs	CUEs	DIF of CUEs
X1	[0.0393,0.0635]	X8.X5	[0.0147,0.0281]	X10.X15	[0.0227,0.0432]
X2	[0.0553,0.0873]	X8.X7	[0.0384,0.0732]	X12.X11	[0.0255,0.0488]
X3	[0.4428,0.6954]	X8.X9	[0.0167,0.0321]	X12.X13	[0.0725,0.1383]
X4.X5	[0.0129,0.0246]	X10.X11	[0.0090,0.0172]	X14.X11	[0.0227,0.0432]
X4.X7	[0.0336,0.0642]	X10.X13	[0.0255,0.0488]	X14.X13	[0.0644,0.1224]
X4.X9	[0.0147,0.0281]	X12.X15	[0.0644,0.1224]	X14.X15	[0.0572,0.1084]

Table 9. EDC of Different diagnostic algorithms

Diagnostic algorithms	EDC
A diagnostic method proposed by Assaf [1]	7.638
A diagnostic method without sensors information	7.211
A diagnostic method with incorporating sensors information	5.519

expressed in interval numbers. Furthermore, qualitative analysis of a DFT is to generate the characteristic function via a zero-suppressed binary decision diagram, while quantitative analysis is to calculate some importance measures by converting a DFT into a DEN. In addition, these reliability results are updated according to the evidence information from sensors and used to design a novel algorithm to improve the diagnosis efficiency. Finally, a real example is given to demonstrate the feasibility and efficiency of the proposed method. This method takes full advantages of both DFT for modelling and DEN for the uncertainty inference, which is especially suitable to diagnose complex systems.

In the future work, we will focus on how the reliability of sensors influences the diagnosis efficiency.

**Acknowledgement**

This research was funded by China Postdoctoral Science Foundation [2015M580568], National Natural Science Foundation of China [71461021] and Graduate Innovation Foundation of Nanchang University [CX2017185].

**References**

1. Assaf T, Dugan J B. Design for diagnosis using a diagnostic evaluation measure. Instrumentation & Measurement Magazine 2006; 9(4): 37-43, <https://doi.org/10.1109/MIM.2006.1664040>.
2. Assaf T, Dugan J B. Diagnosis based on reliability analysis using monitors and sensors. Reliability Engineering & System Safety 2008;

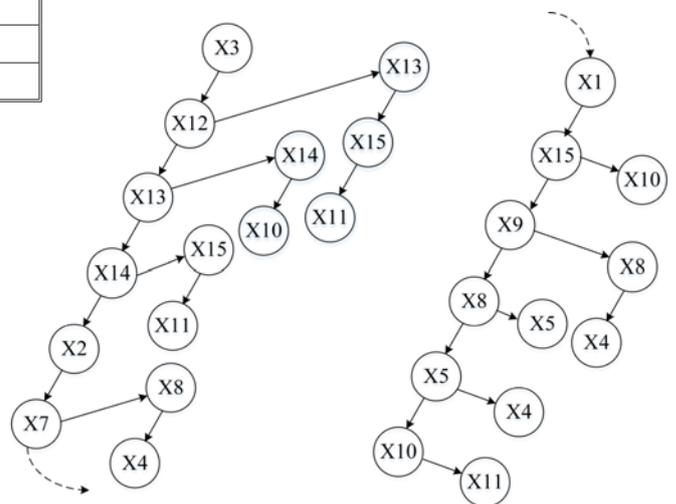


Fig. 5. A corresponding DDT which incorporates sensors information into diagnosis process.

- 93(4): 509-521, <https://doi.org/10.1016/j.res.2006.10.024>.
3. Cai B, Liu H, Xie M. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. *Mechanical Systems and Signal Processing* 2016; 80: 31-44, <https://doi.org/10.1016/j.ymssp.2016.04.019>.
  4. Cao Y, Li T, Wang J, Xie R, Wang X. Fuzzy dynamic fault tree analysis for electro-mechanical actuator based on algebraic model with common-cause failures. *Automatic Control and Computer Sciences* 2016; 50(2): 80-90, <https://doi.org/10.3103/S0146411616020024>.
  5. Chen Y, Zhen Z, Yu H, Xu J. Application of Fault Tree Analysis and Fuzzy Neural Networks to Fault Diagnosis in the Internet of Things (IoT) for Aquaculture. *Sensors* 2017; 17(1): 1-15, <https://doi.org/10.3390/s17010153>.
  6. Chiremsel Z, Said R N, Chiremsel R. Probabilistic Fault Diagnosis of Safety Instrumented Systems based on Fault Tree Analysis and Bayesian Network. *Journal of Failure Analysis & Prevention* 2017; 16(5): 747-760, <https://doi.org/10.1007/s11668-016-0140-z>.
  7. Doguc O, Ramirez-Marquez J E. Using Bayesian approach for sensitivity analysis and fault diagnosis in complex systems. *Journal of Integrated Design and Process Science* 2009; 13(1): 33-48.
  8. Duan R X, Dong D C, Zhao S M. Information Fusion method for system fault diagnosis based on dynamic fault tree analysis. *Journal of Tongji University (Natural Science)* 2011; 39(11): 1699-1704.
  9. Duan R X, Fan J H. A novel diagnosis strategy for braking system of urban rail transit based on fuzzy set and dynamic fault tree. *ICIC Express Letters* 2014; 8(7): 1809-1814.
  10. Duan R X, Zhou H L. Diagnosis strategy for micro-computer controlled straight electro-pneumatic braking system using fuzzy set and dynamic fault tree. *Eksplatacja i Niezawodnosc - Maintenance and Reliability* 2014; 16(2): 217-223.
  11. Duan R X, Hu L F and Lin Y N. Fault diagnosis for complex systems based on dynamic evidential network and multi-attribute decision making with interval numbers. *Eksplatacja i Niezawodnosc - Maintenance and Reliability* 2017; 19(4): 580-589, <https://doi.org/10.17531/ein.2017.4.12>.
  12. Huang H Z, Li Y F, Sun J, Yang Y J. Fuzzy Dynamic Fault Tree Analysis for the Solar Array Drive Assembly. *Journal of Mechanical Engineering* 2013; 49(19): 70-76, <https://doi.org/10.3901/JME.2013.19.070>.
  13. Hui K H, Lim M H, Leong M S, Al-Obaidi S M. Dempster-Shafer evidence theory for multi-bearing faults diagnosis. *Engineering Applications of Artificial Intelligence* 2017; 57: 160-170, <https://doi.org/10.1016/j.engappai.2016.10.017>.
  14. Jiang C, Zhang Z, Han X and Liu J. A novel evidence-theory-based reliability analysis method for structures with epistemic uncertainty. *Computers & Structures* 2013; 129: 1-12, <https://doi.org/10.1016/j.compstruc.2013.08.007>.
  15. Kabir S, Walker M, Papadopoulos Y, Rude E, Securius P. Fuzzy temporal fault tree analysis of dynamic systems. *International Journal of Approximate Reasoning* 2016; 77: 20-37, <https://doi.org/10.1016/j.ijar.2016.05.006>.
  16. Khakzad N, Khan F, Amyotte P. Risk-based design of process systems using discrete-time Bayesian networks. *Reliability Engineering and System Safety* 2013; 109: 5-17, <https://doi.org/10.1016/j.res.2012.07.009>.
  17. Li Y F, Huang H Z, Liu Y, Xiao N C, Li H Q. A new fault tree analysis method: fuzzy dynamic fault tree analysis. *Eksplatacja i Niezawodnosc - Maintenance and Reliability* 2012; 14(3): 208-214.
  18. Liu Z, Li Y F, He L P, Yang Y Z, Huang H Z. A new fault tree analysis approach based on imprecise reliability model. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 2014; 228(4): 371-381, <https://doi.org/10.1177/1748006X14520824>.
  19. Mi J, Li Y., Yang Y J, Peng W, Huang H Z. Reliability assessment of complex electromechanical systems under epistemic uncertainty. *Reliability Engineering & System Safety* 2016; 152: 1-15, <https://doi.org/10.1016/j.res.2016.02.003>.
  20. Philippe W, Christophe S. Dynamic evidential networks in system reliability analysis: A Dempster-Shafer Approach. *Proceedings of the 16th Mediterranean Conference on Control & Automation, Ajaccio, France 2008*; 603-608.
  21. Şahin R, Liu P. Possibility-induced simplified neutrosophic aggregation operators and their application to multi-criteria group decision-making. *Journal of Experimental & Theoretical Artificial Intelligence* 2017; 29(4): 769-785, <https://doi.org/10.1080/0952813X.2016.1259266>.
  22. Sun H L, Yao W X. Comments on methods for ranking interval numbers. *Journal of Systems Engineering* 2010; 25(3): 18-26.
  23. Tao J Y, Dong D C. An improved method for system fault diagnosis using fault tree analysis. *Journal of Harbin Institute of Technology* 2010; 42(1): 143-147.
  24. Tang Z, Dugan J B. Minimal cut set / sequence generation for dynamic fault trees. *Proceedings of Annual Symposium Reliability and Maintainability, Los Angeles, CA, USA, 2004*; 207-213.
  25. Wang D, Zhang Y, Jia X, Jiang P, Guo B. Handling Uncertainties in Fault Tree Analysis by a Hybrid Probabilistic-Possibilistic Framework. *Quality and Reliability Engineering International* 2016; 32(3): 1137-1148, <https://doi.org/10.1002/qre.1821>.
  26. Yuge T, Yanagi S. Dynamic fault tree analysis using Bayesian networks and sequence probabilities. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* 2013; 96(5): 953-962, <https://doi.org/10.1587/transfun.E96.A.953>.
  27. Zhang R, Zhang L, Wang N, Wang W. Reliability evaluation of a multi-state system based on interval-valued triangular fuzzy Bayesian networks. *International Journal of System Assurance Engineering and Management* 2016; 7(1), 16-24, <https://doi.org/10.1007/s13198-015-0335-9>.
  28. Zhang X, Mahadevan S, Deng X. Reliability analysis with linguistic data: An evidential network approach. *Reliability Engineering & System Safety* 2017; 162: 111-121, <https://doi.org/10.1016/j.res.2017.01.009>.

---

**Rongxing DUAN**

**Yanni LIN**

**Yining ZENG**

School of Information Engineering

Nanchang University, 999 Xuefu Rd., Honggutan new district

Nanchang, Jiangxi, China

E-mail: [duanrongxing@ncu.edu.cn](mailto:duanrongxing@ncu.edu.cn), [Nikkilin@yeah.net](mailto:Nikkilin@yeah.net), [1244942196@qq.com](mailto:1244942196@qq.com)

---